Discrepancy between I/B/E/S Actual EPS and Analyst EPS

Lawrence D. Brown
Georgia State University
School of Accountancy
P. O. Box 4050 Atlanta, GA 30302-4050
404-413-7205
accldb@langate.gsu.edu

Stephannie Larocque
University of Notre Dame
398 Mendoza College of Business
Notre Dame, IN 46556
574-631-6136
larocque.1@nd.edu

September 27, 2011

We thank Brad Badertscher, Jeff Burks, William Buslepp, Gus De Franco, Peter Easton, Tom Frecka, Yu Gao, Leo Guo, Mo Khan, Jevons Lee, Greg McPhee, Jeff Miller, Jom Ruangprapun, Pervin Shroff, and Ling Zhou as well as seminar participants at the 2011 AAA Annual Meeting (Denver), Tulane University, University of Minnesota, and University of Notre Dame for their helpful comments and suggestions. We gratefully acknowledge Thomson Reuters for providing I/B/E/S analyst earnings forecast data. Any errors remain our responsibility.

Discrepancy between I/B/E/S Actual EPS and Analyst EPS

Abstract

I/B/E/S includes actual earnings realizations to accompany the earnings-per-share (EPS) forecasts provided by its sell-side analyst contributors. These actuals are available at the firm-year (or firm-quarter) level—they do not vary by analyst. Analysts, however, may provide EPS forecasts to I/B/E/S on differing bases (e.g., they may exclude or include extraordinary items or special items). We investigate the potential discrepancy between the Q1 (Q2, Q3) "actual" EPS furnished by I/B/E/S and the inferred Q1 (Q2, Q3) EPS from a particular analyst's fiscal-year forecast and her forecast for the remainder of the fiscal year– what we refer to as the "inferred" Q1 (Q2, Q3) EPS. We find that the inferred Q1 EPS differs from the I/B/E/S actual Q1 EPS over a third of the time, even when only one analyst follows the firm. This data discrepancy varies systematically at the analyst, firm, industry, and year level, and it is more evident for analysts from small brokerage houses, firms followed by many analysts, firms reporting non-operating or extraordinary items, firms in the transportation and utility industries, and in more recent years. It also occurs more often for Q2 than for Q1, and for Q3 than for Q2. This data discrepancy introduces non-random measurement error into variables that use I/B/E/S actual EPS. We illustrate four adverse statistically and economically significant consequences of this data discrepancy: two pertaining to analysts (less accurate earnings forecasts and smaller earnings revision coefficients) and two pertaining to firms (greater analyst forecast dispersion and smaller market reactions to I/B/E/S earnings surprises). Researchers who use I/B/E/S actual EPS should control for this data discrepancy when either the variables they are interested in include I/B/E/S actual EPS or they include other variables which are correlated with the non-random measurement error we document. We provide two ways for researchers to mitigate the problems we highlight: (1) use an indicator variable approach to control for the measurement error, and (2) incorporate a proxy for the magnitude of the measurement error into their models.

Keywords: I/B/E/S actual EPS, analyst inferred actual EPS, discrepancy, accuracy, revisions, dispersion, surprises.

Discrepancy between I/B/E/S Actual EPS and Analyst EPS

1. Introduction

Many studies use I/B/E/S actual earnings data when they examine analyst forecast accuracy, analyst forecast revision, analyst forecast dispersion, and the market reaction to analyst-based earnings surprises.[1] Other studies control for these variables when they investigate other variables of interest. Researchers conducting these studies act as if the I/B/E/S actuals reliably represent what individual analysts consider to be the actual earnings numbers. If I/B/E/S actuals do not reliably represent analysts' actuals, studies ignoring the discrepancy between the I/B/E/S actual EPS and the analyst EPS introduce measurement error into their analyses, reducing the power of tests of studies that focus on variables incorporating I/B/E/S actual EPS, and reducing the reliability of studies that focus on other factors but which control for variables incorporating I/B/E/S actuals. In the first case, "no-results" studies which were not published may have been published had they obtained results based on more powerful tests. In the second case, studies provide unreliable coefficient estimates of their variables of interest. In both cases, studies addressing the data discrepancy issue would provide more reliable inferences.

In order to elaborate on what we mean by a discrepancy between the I/B/E/S actual EPS and the analyst inferred EPS, we offer the following example. Firm j reports its first quarter EPS at 10am on April 12th. I/B/E/S reports the actual first quarter EPS for firm j as $0.25 at 11am on April 12th. Analysts A and B make forecasts of firm j's second, third, and fourth quarter earnings its annual earnings at noon on April 14th. More specifically,

[1] See Ramnath, Rock, and Shane (2008) for summaries of studies published after 1990, and Brown (2007) for abstracts of articles using analyst earnings data. For a recent article in each area, see Ertimur, Sunder, and Sunder (2007) for accuracy, Feng and McVay (2010) for revision, Barron, Byard, and Kim (2002) for dispersion, and Hugon and Muslu (2010) for market reaction to I/B/E/S earnings surprise.

(1) Both analysts forecast EPS of \$0.25 for Q2, \$0.25 for Q3, \$0.25 for Q4.

(2) Analyst A forecasts \$1.00 for the full fiscal year.

(3) Analyst B forecasts EPS of \$1.05 for the full fiscal year.

To determine whether a data discrepancy exists between the I/B/E/S Q1 actual EPS and the analyst's inferred Q1 EPS, we compare the analyst's forecast for the fiscal year with the sum of the I/B/E/S Q1 EPS and the analyst's forecast for the remaining quarters of the year. In the case of analyst A, the numbers do add up (i.e., the \$1.00 forecast for the full fiscal year equals the sum of the I/B/E/S Q1 actual EPS of \$0.25 and the analyst's forecast of \$0.75 for the remainder of the year). In the case of analyst B, the numbers do not add up (i.e., the \$1.05 forecast for the full fiscal year is five cents greater than the sum of the I/B/E/S Q1 actual EPS of \$0.25 and the analyst's forecast of \$0.75 for the remainder of the year). In other words, analyst A includes \$0.25 of Q1 EPS in her fiscal-year forecast, identical to the I/B/E/S Q1 actual EPS. We refer to this \$0.25 as analyst A's inferred Q1 EPS. In contrast, analyst B includes \$0.30 of Q1 EPS in his fiscal-year forecast, five cents more than the I/B/E/S Q1 actual EPS. We refer to this \$0.30 as analyst B's inferred Q1 EPS. We say that a data discrepancy does not exist for analyst A but that it does exist for analyst B.[2]

We find that this data discrepancy, defined as cases where the I/B/E/S actual EPS differs from the analyst's inferred EPS by at least one penny, occur 35%, 45%, and 52% of the time for the first (Q1), second (Q2), and third (Q3) interim quarters, respectively. We find that data discrepancy has analyst, firm, industry, and year effects, and that it occurs

[2] Some the discrepancy is undoubtedly due to the analyst and some of it to I/B/E/S. Some of the discrepancy may be intentional (strategic) and some of it unintentional (information processing difficulties). It beyond the scope of our study to determine how much of the discrepancy is due to the analyst, how much to I/B/E/S, how much is intentional, and how much is unintentional. Our goal is to introduce the problem, determine some of the factors it is associated with, highlight its importance to researchers, and provide a means to mitigate its consequences

more often for analysts employed by smaller brokerage houses,[3] for firms with non-operating or extraordinary items on their income statements, for firms followed by more analysts, and for firms in the transportation and utility industries. We show that the data discrepancy increases over the time period of our study. We document four statistically and economically significant adverse consequences of this data discrepancy, two pertaining to analysts (reduced forecast accuracy and smaller earnings revision coefficients) and two pertaining to firms (greater analyst forecast dispersion and smaller market reactions to I/B/E/S earnings surprises).

We proceed as follows. We present our research questions in section 2. We discuss our sample selection procedures and show the prevalence of data discrepancy in section 3. We examine factors associated with the data discrepancy in sections 4 and 5, respectively. We provide robustness tests in section 6 and conclusions in section 7. The Appendix contains definitions of all the variables we use in our study.

2. Research questions

In order to determine whether a discrepancy exists between the Q1 I/B/E/S actual EPS and the analyst's inferred Q1 EPS, we compare the Q1 I/B/E/S actual EPS with the inferred Q1 analyst EPS for all analysts in the I/B/E/S database possessing the requisite information. Prior to the Q1 earnings announcement, an analyst forecasts earnings of both Q1 and FY (the fiscal year). After the Q1 earnings are announced, the same analyst forecasts Q2, Q3, Q4, and FY. We refer to the analyst's forecasts prior to (after) the Q1 earnings

[3] This result is consistent with the notion that I/B/E/S pays more attention to opinions of large brokerage houses. Ljungqvist, Malloy, and Marston (2009), who study analyst recommendations data, reach a similar conclusion.

announcement as her first (second) forecast, and we measure her inferred Q1 EPS by subtracting the summation of her second forecast of Q2, Q3, and Q4 from her second forecast of FY. Our first task is to determine how often the data discrepancy occurs since the frequency of its occurrence is directly related to its potential impact on studies that do not attempt to mitigate its effects. Our first research question (RQ1) is:

RQ1: How often does a discrepancy exist between the I/B/E/S Q1 actual EPS and the analyst's Q1 inferred EPS?

After providing evidence that this data discrepancy is pervasive in our sample, we investigate whether it is random or systematic. If it is systematic, it has greater potential impact on studies whose data possess the systematic factors that we document. We consider four possible factors that may be associated with this data discrepancy: analyst effects, firm effects, industry effects, and year effects. Our second research question is:

RQ2: Is the data discrepancy we document related to the following systematic factors: analyst effects, firm effects, industry effects, or year effects?

After finding evidence that the data discrepancy we introduce is associated with all four of these fixed effects, we examine its consequences for four common types of studies using I/B/E/S actual EPS data.[4] The importance to researchers of the data discrepancy we introduce depends upon the magnitude of its effects on the coefficients that researchers seek to determine.[5] Our third research question is:

[4] We do not query whether any effects exist because the measurement error caused by data discrepancy guarantees that the effects exist (i.e., adding measurement error to an independent variable mitigates the magnitude of the variable's coefficient estimate). Our task is to determine the economic significance of the measurement error caused by the inconsistency.

[5] For simplicity, we only consider the direct effects of the data discrepancy. We do not consider its effects on studies whose focus is on variables that do not incorporate I/B/E/S actual EPS but which include these data in their control variables. Thus, our findings probably understate the potential problems to researchers who do not address the data discrepancy we introduce.

RQ3: What are the consequences to researchers of not controlling for the data discrepancy we document for studies of forecast accuracy, forecast revisions, forecast dispersion, and the capital market reaction to earnings surprises?

We find that the consequences of not controlling for the data discrepancy we document are both statistically significant and economically important in that the magnitudes of its effects on the coefficients of all the variables using I/B/E/S actuals are material. We provide two robustness checks. First, we examine if our results pertain to subsequent interim quarters. We find that they do, especially following the third quarter, the time frame for many studies which use analyst data shortly before annual earnings announcements. Second, we examine whether our results are robust to using a proxy for the magnitude of the data discrepancy we introduce. We find that they are.

3. Sample selection procedures and evidence regarding prevalence of data discrepancy

We extract from the I/B/E/S earnings detail files U.S. firms with available reporting dates for fiscal year *t-1* (FY_{t-1}) and the first and second quarters of fiscal year *t* ($Q1_t$ and $Q2_t$) for the 13 years, 1996 to 2008. We require reporting dates for FY_{t-1}, $Q1_t$, and $Q2_t$ because we examine analyst forecasts made: (1) after the year *t-1* report but before the first quarter of year *t* report; and (2) after the first quarter of year *t* report but before the second quarter of year *t* report. Of the 56,366 firm-years with available data for FY_{t-1}, $Q1_t$, and $Q2_t$, we retain 47,615 firm-years for which we can obtain actual earnings (EPS_t) from I/B/E/S for FY_t and $Q1_t$.[6] We retain 46,037 firm-years with share prices as of the end of year *t-1* available from CRSP.

[6] While we require year *t-1*'s reporting date we do not require its actual earnings.

For these 46,037 firm-years, we search the I/B/E/S detail file for analysts who, after the release of $Q1_t$ earnings, issued EPS forecasts for $Q2_t$, $Q3_t$, $Q4_t$, and FY_t on the same day. We limit the sample to the 31,818 analyst-firm-years who issued a $Q1_t$ forecast after the release of FY_{t-1} earnings but before the release of $Q1_t$ earnings. We use these 31,818 analyst-firm-years to examine the impact of data discrepancy on analyst forecast accuracy of $Q1_t$[7]. 36% (64%) of analyst-firm-years are followed by only (more than) one analyst. We require that multiple analysts follow a firm when we examine the association of data discrepancy with analyst effects and analyst earnings forecast dispersion.

We form four sub-samples of the 31,818 analyst-firm-years. We use: (1) 23,921 observations with FY_t forecasts made after the release of FY_{t-1} earnings but before the release of $Q1_t$ earnings to examine the impact of data discrepancy on analyst forecast accuracy of FY_t; (2) 20,968 observations with $Q2_t$, $Q3_t$, and $Q4_t$ forecasts made after release of FY_{t-1} earnings but before release of $Q1_t$ earnings to examine the impact of data discrepancy on analyst earnings forecast revisions; (3) 7,108 (5,078) firm-years with multiple analyst following to examine the impact of data discrepancy on analyst forecast dispersion in $Q1_t$ (FY_t) earnings forecasts; and (4) 18,465 observations with non-missing CRSP stock return data to examine the impact of data discrepancy on the capital market reaction to I/B/E/S earnings surprises. Table 1 outlines our sample selection method.

Table 1

We examine the analyst's annual earnings forecast made after Q1 earnings is known. If the analyst's annual earnings forecast made after Q1 earnings is within a penny of

[7] These data are also the starting point for our analysis of the sources of data inconsistency. As discussed in the next section, additional data restrictions apply to the analysis of each source.

the sum of the I/B/E/S Q1 actual and the analyst's forecast for the remainder of the year, we say that the analyst's inferred EPS conforms to the I/B/E/S actual Q1 EPS, and we set *DISCREP* (discrepancy) equal to 0. If the analyst's annual earnings forecast made after Q1 earnings is *not* within a penny of the sum of the I/B/E/S Q1 actual plus the analyst's forecast for the remainder of the year, we say that the analyst's inferred EPS does not conform to the I/B/E/S actual Q1 EPS, and we set *DISCREP* equal to 1.[8] Appendix 2 illustrates.

For 20,812 observations or 65% of the analyst-firm-years, *DISCREP* equals 0. Thus, 11,006 or 35% of the analyst-firm-years have *DISCREP* = 1. Figure 1 presents the distribution of the difference between the analysts' inferred Q1 EPS and the I/B/E/S Q1 actual. The median difference is 0.000 and the mean of -0.003 is insignificantly different from zero (t-statistic = 0.08).

Figure 1

4. Factors associated with data discrepancy

In order to determine if the data discrepancy is systematic or random, we estimate the following fixed effects model using the logistic regression:

$$DISCREP_{ijkt} = \mu_i + \delta_j + \lambda_k + \gamma_t + \varepsilon_{ijkt} \quad (1)$$

The four factors we consider are (1) analyst effects (μ_i), (2) firm effects (δ_j), (3) industry effects (λ_k), and (4) year (γ_t) effects. The dependent variable, *DISCREP*, equals one when analyst *i* following firm *j* in I/B/E/S industry *k* in year *t* has an annual forecast for year *t* made following the release of Q1 EPS that is not within one penny of the sum of the I/B/E/S Q1

[8] When we use $DISCREP_{ijt}$ in our accuracy and revision analyses, the term refers to analyst-firm-years. When we use $DISCREP_{jt}$ in our dispersion and market reaction to earnings surprise analyses, the term refers to firm-years.

actual EPS and the analyst's forecast of the remainder of the year and zero otherwise. The model residual is represented by ε_{ijkt}. Following O'Brien (1990), we estimate four separate fixed effects models where each includes one fixed effect. Following Hamilton (1992), we evaluate the log-likelihood of each logistic model to determine if the fixed effect helps to explain the probability that the discrepancy occurs.

Our estimations require multiple observations of: (1) analysts for the analyst model; (2) firms for the firm model; (3) industries for the industry model; and (4) years for the year model. Table 2 presents the results. When we include only analyst fixed effects, panel A of Table 2 indicates that the log-likelihood ratio is -8,466.0[9]. When we include only firm fixed effects, the log-likelihood ratio is -11,857.3. When we include only industry fixed effects, the log-likelihood ratio is -18,162.5. When we include only year fixed effects, the log-likelihood ratio is -18,069.8. When we include analyst, firm, industry, and year effects, the log-likelihood ratio is -3,436.9. In sum, the data discrepancy is systematic, being associated with analyst, firm, industry, and year.

Table 2

To learn more about these four effects, we investigate each one separately. Panel B of Table 2 presents the association of *DISCREP* with five analyst-level variables commonly used in the literature (Mikhail, Walther, and Willis 1997; Clement 1999; Jacob, Lys, and Neale 1999; Clement and Tse 2005): (1) firm experience (*FEXP*), (2) forecast frequency (*FREQ*), (3) number of firms covered (*NFIRMS*), (4) brokerage size (*BSIZE*), and (5) forecast horizon (*HORIZON*). Our results reveal that the data discrepancy is less prevalent for

[9] The χ^2 for each model is significant at the 0.00 level.

analysts who work for large brokerage houses, and it is unrelated to the other factors we consider. It is plausible that I/B/E/S "listens more" to the larger brokerage houses when it selects the Q1 actual to put in its database.

Panel C examines the first firm effect, the number of analysts following the firm. The frequency of data discrepancy increases monotonically as the number of analysts following the firm increases, but it still occurs nearly one-third of the time when only one analyst (with the necessary data available for our study) follows a firm. Thus, the frequency of measurement errors we document is substantial, and it is not solely attributable to the fact that I/B/E/S actuals represent the majority opinion of analysts.[10]

Panel D examines the second firm effect: the presence or absence of non-operating or extraordinary items. We find that firms reporting Q1 EPS containing non-operating or extraordinary items are more likely to display this data discrepancy. Our results pertain to all 11 I/B/E/S industries, and are significant at the 6% level or better in eight of them.

Panel E provides pooled descriptive statistics on the percent of data discrepancy for each of the 11 industries: Basic Industries (Basic), Capital Goods (Capital), Consumer Durables (ConsDur), Consumer Non-Durables (ConsNonDur), Consumer Services (ConsSvc), Energy (Energy), Finance (Finance), Health Care (Health), Technology (Technol), Transportation (Transp), and Public Utilities (Utility). Data discrepancy ranges from less than 30 percent for the capital goods and consumer non-durables industries to more than 42 percent for the transportation and utility industries.

[10] According to I/B/E/S, "when a company reports their earnings, the data is evaluated by a Market Specialist to determine if any Extraordinary or Non-Extraordinary Items (charges or gains) have been recorded by the company during the period… If one or more items have been recorded during the period, actuals will be entered based upon the estimates majority basis at the time of reporting."

Panel F examines year effects by estimating the following specification for each of the 11 industries separately and for the full sample:

$$Pr(DISCREP_{ijt} = 1) \quad = a + b * Year_{ijt} + e_{ijt} \tag{2}$$

To facilitate interpretation of our results, we set the first year of our sample (1996) equal to one, the second year equal to two, and so on. Our results reveal that data discrepancy in the I/B/E/S earnings data base has significantly increased over time for ten of the 11 industries and the full sample.

In sum, the likelihood of data discrepancy is higher for: (1) analysts employed by small brokerage houses; (2) firms with non-operating or extraordinary items; (3) firms with greater analyst following; (4) firms in the transportation and utility industries; and (5) in more recent years. Un-tabulated results reveal that discrepancy between I/B/E/S actual EPS and analyst EPS equals 50% in the last three years of our sample (2006-2008).

5. Consequences of the data discrepancy

We provide univariate results for the full sample, the *DISCREP = 0* and *DISCREP = 1* sub-samples, and the difference between the sub-samples. In multivariate analyses, we document the severity of the consequences of the data discrepancy for studies of forecast accuracy, forecast revision, forecast dispersion, and market reaction to I/B/E/S-derived earnings surprise. To mitigate the impact of outliers, we winsorize the distributions of accuracy, revisions, and earnings forecasts in the dispersion analysis, and the distribution of I/B/E/S-derived earnings surprises in the top and bottom 1% of observations.

5.1 Earnings forecast accuracy

Table 3 presents separate results for modeling the accuracy of analysts' forecasts of both first-quarter EPS and annual EPS. For brevity, we only discuss first-quarter results in the text, but results for annual forecasts are qualitatively similar. *Accuracy* is defined as the absolute value of the difference between I/B/E/S actual earnings and the analyst's last EPS estimate before the release of Q1 EPS, scaled by lagged stock price so that a smaller value indicates greater accuracy.[11] We expect analysts whose inferred Q1 EPS differs from the I/B/E/S Q1 actual to make less accurate earnings forecasts because accuracy is based on the I/B/E/S actual EPS rather than the analyst's inferred EPS. Thus, she is evaluated against a different target than she was aiming at. Panel A presents pooled accuracy results for the full sample. Earnings forecast accuracy averages 0.0048 across all analyst-firm-years, and it is 0.0039 and 0.0066 for *DISCREP* = 0 and *DISCREP* = 1 analyst-firm-years, respectively. The order of magnitude of unsigned error for *DISCREP = 1* is nearly 70% larger than that for *DISCREP = 0*, a difference which is both economically and statistically significant (t-statistic = 18.19).

Panel B provides results of a multivariate analysis to examine the relation between forecast accuracy and data discrepancy. We estimate the following model:

$$\begin{aligned} \mathit{Accuracy}_{\mathrm{ijt}} &= \alpha_0 + \alpha_1 \mathit{DISCREP}_{ijt} + \alpha_2 \mathit{FEXP}_{ijt} + \alpha_3 \mathit{FREQ}_{ijt} + \alpha_4 \mathit{NFIRMS}_{ijt} + \alpha_5 \mathit{BSIZE}_{ijt} \\ &\quad + \alpha_6 \mathit{HORIZON}_{ijt} + \alpha_7 \mathit{NONOP}_{jt} + \Sigma\text{Analyst} + \Sigma\text{Firm} + \Sigma\text{Industry} + \Sigma\text{Year} + e_{\mathrm{ijt}} \end{aligned} \tag{3}$$

We control for five analyst-firm-year effects: firm experience (*FEXP*), forecast frequency (*FREQ*), number of firms followed *(NFIRMS*), brokerage size (*BSIZE*), and forecast

[11] We use the last forecast in lieu of the consensus because it more closely represents market expectations (Brown and Kim 1991).

horizon (*HORIZON*). We also control for instances where firms have non-operating or extraordinary items on their income statements (*NONOP*) along with analyst, firm, industry, and year fixed effects.[12] Based on prior research (Mikhail et al. 1997; Clement 1999; Jacob et al. 1999; Clement and Tse 1995), we expect analysts who are more experienced, forecast more often, and work for larger brokerage houses to be more accurate. We expect analysts who follow more firms, make forecasts over longer time horizons, and forecast firms reporting non-operating or extraordinary items to be less accurate. The coefficients on five of the control variables have their expected signs.[13] Analysts are less accurate if they make forecasts over longer horizons, follow more firms, or forecast earnings with non-operating or extraordinary items.[14] Analysts who forecast more often and who are employed at large brokerage houses are more accurate, but these effects are not significant.

As in the univariate analyses, we find that data discrepancy leads to significantly less accurate forecasts. To assess the economic significance of our results, we compare the coefficient on *DISCREP* (α_1) with the intercept (α_0), which represents the accuracy of analysts whose inferred Q1 EPS are within one penny of the I/B/E/S Q1 actual EPS (after controlling for the other factors in the model). Analysts' one-quarter-ahead forecasts are 80% less accurate (0.0004/0.0005) when *DISCREP* = 1 rather than 0.[15] As evidence of the statistical importance of our evidence, *DISCREP* has a larger t-value than all five analyst-firm-year effects we include in our model. Its t-value of 3.12 is only slightly less than the t-value

[12] *NONOP* = 1 when the absolute value of the difference between Q1 EPS Compustat operating earnings and Compustat earnings after extraordinary items for year *t* exceeds one penny. For convenience we drop the i, j and t subscripts in the text but we retain them in the equations and in the tables. We do not include a subscript for industry because it is redundant (i.e., once the firm-year is known, the industry is known).

[13] While unexpected, our results are similar to those of Jacob et al. (1999), who do not find evidence that more experienced analysts are more accurate.

[14] The coefficients on *NFIRMS* and *NONOP* are insignificant in the multivariate analysis of annual forecasts.

[15] Analysts' annual earnings forecasts are 82% less accurate (0.0023/0.0114) when analysts' actuals are inconsistent with I/B/E/S actuals.

of *NONOP* of 3.27. The magnitude of *DISCREP*'s coefficient of 0.0004 is identical to that of *NONOP*, the other indicator variable.

Table 3

5.2 Earnings forecast revisions

Table 4 presents results relating to the revision of analysts' EPS forecasts for the last nine months of the fiscal year. Panel A estimates the following equation:

$$REV\_9M_{ijt} = a + b * SURP_{ijt} + e_{ijt} \tag{4}$$

The dependent variable (*REV_9M*) represents the revision of analyst *i*'s EPS estimate for firm *j* for Q2 through Q4 of year *t*, defined as the difference between the analyst's first forecast of Q2 through Q4 issued after the release of Q1 earnings and before the release of Q2 earnings and the same analyst's last forecast of Q2 through Q4 earnings issued after the release of year *t-1* earnings and before the release of Q1 earnings, scaled by lagged stock price. *SURP* is defined as the difference between the I/B/E/S actual Q1 earnings and the last Q1 forecast issued by analyst *i* before the release of Q1 earnings, scaled by lagged stock price. We expect the beta coefficient for the sub-sample without the data discrepancy to be higher than the beta coefficient for the sub-sample with the data discrepancy because measurement error in *SURP* drives the beta coefficient towards zero.

Table 4 Panel A presents pooled revision results for the sub-sample of 20,968 analyst-firm-years with the same analyst's earnings forecasts of Q1, Q2, Q3, and Q4 EPS made on the same day before the release of Q1 earnings. The revision slope coefficient of 1.09 for all analyst-firm-years reveals that for every $1.00 of *SURP* for Q1, on average, analysts revise their EPS forecasts of the remainder of the year in the same direction of the surprise by \$1.09 (t-statistic = 45.32). As in Table 3, we partition our sample into *DISCREP = 0* and *DISCREP = 1*. The forecast revision coefficient for the first sub-sample is significantly larger than the forecast revision coefficient for the latter sub-sample (1.37 versus 0.87, t-statistic of the difference = 10.35). In other words, when the inferred analyst Q1 EPS is equal to the I/B/E/S Q1 actual EPS, for every \$1.00 of *SURP,* on average, they revise their EPS forecast for the remainder of the year in the same direction of the surprise by \$1.37 (t-statistic = 42.11). In contrast, when the inferred analyst Q1 EPS differs from the I/B/E/S Q1 actual EPS, for every \$1.00 of *SURP,* on average, analysts revise their EPS forecast for the remainder of the year in the same direction of the surprise by only \$0.87 (t-statistic = 23.02). The order of magnitude of the revision coefficient for *DISCREP = 0* is more than 57% higher than that for *DISCREP = 1*, a difference which is both economically and statistically significant (t-statistic = 10.35).[16]

Panel B provides results which control for *NONOP* along with analyst, firm, industry, and year fixed effects. More specifically, it estimates the following model:

$$\begin{aligned} REV\_9M_{ijt} &= \alpha_0 + \alpha_1 DISCREP_{ijt} + \alpha_2 SURP_{ijt} + \alpha_3 DISCREP_{ijt} * SURP_{ijt} + \alpha_4 NONOP_{jt} \\ &\quad + \Sigma\text{Analyst} + \Sigma\text{Firm} + \Sigma\text{Industry} + \Sigma\text{Year} + e_{ijt} \end{aligned} \tag{5}$$

We have no priors regarding *DISCREP*'s main effect (α_1). Given that analysts revise their forecasts of future quarters in the direction of their most recent forecast error (Brown and Rozeff 1979), we expect analysts' revisions (α_2) to be positively related to *SURP*. Because *SURP* measures the analyst's *SURP* with error when the analyst's inferred Q1 EPS differs from the I/B/E/S Q1 actual, we expect the revision coefficient on *SURP* to be smaller

[16] The order of magnitude of the unsigned error for the latter group is nearly 50% larger than that for the former group in the univariate analysis of annual forecasts (t-statistic = 12.18).

when *DISCREP = 1* (α_3). As non-operating and extraordinary items are more transitory than operating earnings (Elliott and Hanna 1996), we expect analysts to revise their forecasts of future earnings to a lesser extent conditional on *SURP* when it contains a non-operating or extraordinary item component (α_4).

The main effect of *DISCREP* is small in magnitude and marginal in significance (coefficient = 0.0006, t-value = 1.65). As expected, the coefficient on *SURP* is positive and significant (0.9205, t-value = 26.52). Most importantly for our purposes, the coefficient on *DISCREP*SURP* is negative and significant (-0.2505, t-value = -5.08). Thus, analysts whose inferred Q1 EPS equals the I/B/E/S Q1 actual EPS revise their earnings forecasts for the remainder of the year 37% more than do those analysts whose inferred Q1 EPS does not equal the I/B/E/S Q1 actual EPS. Our results are both economically and statistically significant.[17]

Table 4

5.3 Earnings forecast dispersion

Table 5 presents results for the dispersion of analysts' first-quarter and annual EPS estimates. For brevity, we only discuss the first-quarter results in the text but the annual results are qualitatively similar. We expect mean dispersion to be smaller when all analysts are shooting at the same target than when some analysts are shooting at different targets than other analysts. In Panel A, we present estimates of dispersion as a pooled average of the standard deviation of analysts' EPS estimates for each firm-quarter for firms with multiple analysts following them in a given firm-quarter. Mean dispersion (*STD* (Q1)) for

[17] As expected, *NONOP* has a negative coefficient but it is not significantly different from zero.

these 7,108 firm-years is 0.044. Mean dispersion averages 0.027 for firm-years with *DISCREP* = 0 (all analysts are shooting at the same target), and 0.056 for firm-years with *DISCREP* = 1 (some analysts are shooting at different targets). Thus, mean dispersion for firm-years with data discrepancy is more than double that for firm-years with data discrepancy. [18]

Panel B presents results which control for firms' earnings stability (*STABLE*), *NONOP*, and fixed effects for firm, industry and year. More specifically, we estimate the model:

$$STD\,(Q1)_{jt} = \alpha_0 + \alpha_1 DISCREP_{jt} + \alpha_2 STABLE_{jt} + \alpha_3 NONOP_{jt} + \Sigma\text{Firm} + \Sigma\text{Industry} + \Sigma\text{Year} + e_{jt} \qquad (6)$$

We expect less agreement among analysts when *DISCREP = 1* than when *DISCREP = 0* so we expect the coefficient on *DISCREP* (α_1) to be positive. As there is likely to be more agreement among analysts when earnings are more stable, we expect the coefficient on *STABLE* (α_2) to be negative. As there is likely to be less agreement among analysts when firms report non-operating or extraordinary items, we expect *NONOP* to have a positive coefficient (α_3).

The coefficients on the two control variables have their expected signs and both are significant. More specifically, forecasts are less dispersed when earnings are more stable (α_2 =-0.00004, t-value = -2.56) and forecasts are more dispersed when firms report non-operating or extraordinary items (α_3 =0.0016, t-value = 1.48).[19]

[18] Researchers who use analyst dispersion as a proxy for difference of opinion should be aware that dispersion may reflect both differences of opinion regarding an earnings number defined a certain way (e.g., all analysts' forecasts include restructuring charges) and differences of opinion regarding an earnings number defined in different ways (e.g., some analysts' forecasts exclude restructuring charges while other analysts' forecasts include restructuring charges). Thus, the researcher's proxy variable includes both differences of opinion regarding an earnings number(s) and definitional differences of what earnings number analysts are thinking about.

[19] The control variables are insignificant in the analysis of annual forecasts.

Most importantly for our purposes, dispersion is greater for when *DISCREP* = 1 than when *DISCREP* = 0 (α_2 = 0.0043, t-value = 3.85). To determine the economic significance of our results, we compare the α_1 coefficient with the intercept (α_0) which represents dispersion for firms followed by analysts who all agree with the I/B/E/S Q1 actual. Analyst dispersion is 23% greater (.0043/.0186) when *DISCREP* = 1 for at least one analyst following the firm. Our results are both economically and statistically significant.[20]

Table 5

5.4 Market reaction coefficients

Table 6 provides results from estimating the market reaction to the first quarter I/B/E/S earnings surprise. Panel A presents results based on the following specification:

$$CAR_{jt} = a + b * SURP_{jt} + e_{jt} \tag{7}$$

CAR is the two-day (-1, 0) cumulative return minus the cumulative value-weighted market return related to the announcement of Q1 earnings for firm *j* in year *t*. *SURP* is defined as the I/B/E/S actual Q1 EPS minus the last Q1 estimate issued by an analyst following firm *j* before the release of Q1 earnings, scaled by lagged stock price.[21] We expect the beta coefficient for the sub-sample without the data discrepancy to be higher than the beta coefficient for the sub-sample with the data discrepancy because measurement error in *SURP* drives the beta coefficient towards zero. Using the entire sample for which we can calculate *CAR*, *b* in equation (5) is estimated to be 0.976 (t-statistic = 17.78) conforming to the well-documented evidence that abnormal returns are positively and significantly related

[20] The t-statistic of *DISCREP* used to test RQ3 is 3.85.

[21] The analyst in both this regression and the earnings forecast revision regression (equation 5) may be either a *CON* or an *INCON* analyst. Thus, our results probably understate the effects that we document.

to quarterly earnings surprises (Brown and Kennelly 1972; Foster 1977). As in Tables 3 to 4, we partition our sample into *DISCREP* = 0 and *DISCREP* = 1 sub-samples. Column 2 includes firm-years for which *DISCREP* = 0. Column 3 includes firm-years for which *DISCREP* = 1 for at least one analyst following the firm. The *b* coefficient for the market reaction coefficient for the *DISCREP* = 0 sub-sample is nearly double that of the *DISCREP* = 1 sub-sample (1.364/0.717). This difference is both economically and statistically significant (t-statistic = -5.78).

Panel B provides results which control for *STABLE*, *NONOP*, and fixed effects for firm, industry, and year. More specifically, we estimate the following model:

$$CAR_{jt} = \alpha_0 + \alpha_1 DISCREP_{jt} + \alpha_2 SURP_{jt} + \alpha_3 DISCREP_{jt} * SURP_{jt} + \alpha_4 STABLE_{jt} + \alpha_5 NONOP_{jt} + \Sigma\text{Firm} + \Sigma\text{Industry} + \Sigma\text{Year} + e_{jt} \quad (8)$$

Firms with stable earnings should have a higher valuation multiplier (Beaver 1998) so we expect α_4 to be positive. The market should react less to firms' earnings with non-operating or extraordinary items as their earnings are more transitory (Elliott and Hanna 1996) so we expect α_5 to be negative. Scores of studies have shown that earnings surprises are value-relevant (Ball and Brown 1968; Foster 1977; Beaver et al. 1979) so we expect α_2 to be positive. We have no priors for the main effect of *DISCREP* so we have no expectation as to the sign of its coefficient (α_1), but we do have a prior for the coefficient on our primary variable of interest: the interaction of *DISCREP with SURP* (α_3). The positive relation between market reaction and earnings surprise should be attenuated when *DISCREP* = 1 so we expect α_3 to be negative.

Investors react marginally more positively to I/B/E/S earnings surprises of firms with more stable earnings (α_4 = 0.00003, t-value = 1.50) but they do not react more negatively to

NONOP (α_5 = 0.0010, t-value = 0.80). There is a positive market reaction to I/B/E/S earnings surprises for firm-years when *DISCREP* = 0 (α_2 = 2.1154, t-value = 11.95). Most importantly for our purposes, the positive reaction to earnings surprises is mitigated for firm-years when *DISCREP* = 1 (α_3 = -0.6205, t-value = -2.75). To determine the economic significance of these results we take the ratio of the coefficient of *DISCREP*SURP* to that of *SURP*. The market reaction to the earnings surprise is over 41 percent greater when *DISCREP* = 0 than when *DISCREP* = 1. Thus, the results for our key variable of interest are both economically and statistically significant.

Table 6

6. Robustness tests

6. 1. Second and third quarter earnings

To determine if our results are robust to the other two interim quarters, we repeat our analyses for the intervals:

1. Following the release of $Q2_t$ earnings but before the release of $Q3_t$ earnings; and
2. Following the release of $Q3_t$ earnings but before the release of $Q4_t$ (fiscal year *t*) earnings.

In these two respective intervals, data discrepancy exists 45% and 52% of the time. As the frequency of data discrepancy is 35% for the first quarter, it is evident that data discrepancy increases substantially throughout the fiscal year.

We present analyst-firm-level accuracy and revision results in Table 7 Panel A. The accuracy results for Q2 conform to those for Q1 we reported in panel B of Table 3, while the revision results for both *REV_6M* and *REV_3M* are similar to the revision results for *REV_9M* we reported in panel B of Table 4. Panel B of Table 7 provides results for forecast dispersion and market reaction to I/B/E/S earnings surprises for Q2 and Q3 in a manner similar to what we reported for Q1 in panel B of Tables 5 and 6. Consistent with the dispersion results for Q1, data discrepancy significantly heightens forecast dispersion for Q3, but not for Q2 (see the coefficients and corresponding t-values of *DISCREP*). Consistent with the market reaction to I/B/E/S earnings surprise results for Q1, data discrepancy significantly mitigates the capital market reaction to I/B/E/S earnings surprises for both Q2 and Q3 (see the coefficients and corresponding t-values of *DISCREP*SURP*).

In sum, data discrepancy is more prevalent with I/B/E/S Q3 actuals than with I/B/E/S Q2 actuals and it is more prevalent with I/B/E/S Q2 actuals than with I/B/E/S Q1 actuals. The results we document for the first interim report for both forecast revisions and the capital market reaction to earnings surprises also pertains to the second and third interim reports. The results we document for Q1 for forecast accuracy and forecast dispersion pertain to Q3 but not to Q2. Overall, our results are quite robust to using interim quarters two and three (especially the latter) in lieu of the first interim quarter.

____

Table 7

____

6.2. A continuous measure of data discrepancy

To this point we have measured *DISCREP* as a zero-one indicator variable based on whether there is at least a penny difference between the analyst's inferred EPS and the I/B/E/S actual EPS. This approach has two disadvantages: (1) it ignores the magnitude of the data discrepancy; and (2) the penny cutoff is arbitrary. We repeat our multivariate analyses of accuracy, revision, dispersion, and market reaction for Q1 using the following continuous measure data discrepancy:

$$MDISCREP_{ijt} = \textit{Analyst Inferred EPS } (Q1)_{ijt} - EPS\ (Q1)_{ijt} \quad (9)$$

MDISCREP is a proxy variable for the magnitude of the data discrepancy; *Analyst Inferred EPS* (*Q1)* is the difference between the analyst's forecast of the year minus the forecast of Q2 + Q3 + Q4, conditional on Q1 having been reported; and *EPS* is Q1 earnings as reported by I/B/E/S.

Thus, we employ the following multivariate regression model, which is a modification of equation (3), to examine accuracy:

$$\begin{aligned} Accuracy_{ijt} = {} & \alpha_0 + \alpha_1|MDISCREP_{jt}| + \alpha_2 FEXP_{ijt} + \alpha_3 FREQ_{ijt} + \alpha_4 NFIRMS_{ijt} + \alpha_5 BSIZE_{ijt} \\ & + \alpha_6 HORIZON_{ijt} + \alpha_7 NONOP_{jt} + \Sigma\text{Analyst} + \Sigma\text{Firm} + \Sigma\text{Industry} + \Sigma\text{Year} + e_{ijt} \end{aligned} \quad (10)$$

Equation (10) differs from equation (3) only in that it uses the absolute value of *MDISCREP* in lieu of *DISCREP*. Panel A of Table 8 indicates that the coefficient on |*MDISCREP*| (α_1) is positive and significant (coefficient = 0.0001; t-value = 4.11). Moreover, it has the largest t-value of any of the independent variables, indicating that it provides the most incremental power of any of the independent variables for explaining the variability in analyst forecast accuracy.[22] Equation (10) has more total explanatory power than equation (3) [1.1% versus 1.0%], revealing that the *MDISCREP* model is more powerful than the *DISCREP* indicator variable for the purpose of explaining the variability in analyst forecast accuracy.

We employ the following multivariate regression which is a modification of equation (5), to examine analyst forecast revisions:

[22] Recall that *NONOP* had a higher t-value than *INCON* in the equation (3) estimation.

$$REV\_9M_{ijt} = \alpha_0 + \alpha_1 MDISCREP_{ijt} + \alpha_2 SURP_{ijt} + \alpha_3 NONOP_{jt} + \Sigma\text{Analyst} + \Sigma\text{Firm} + \Sigma\text{Industry} + \Sigma\text{Year} + e_{ijt} \quad (11)$$

Equation (11) differs from equation (5) in two ways: (1) it uses *MDISCREP* in lieu of *DISCREP* and (2) it omits the interaction term of equation (5).[23] While the focal point of equation (5) is on the interaction term between *DISCREP* and *SURP*, the focal point of equation (11) is on *MDISCREP*. We expect the analyst to revise his earnings forecast of the remainder of the year in the direction of his inferred error which is equal to the sum of *MDISCREP* and *SURP*. Thus, we expect the coefficient on *MDISCREP* (α_1) to be positive. Consistent with our expectations, panel A of Table 8 indicates that the coefficient of *MDISCREP* is positive and significant (coefficient = 0.0003, t-value = 2.11). Equation (11) has less total explanatory power than equation (5) [8.1% versus 8.2%], revealing that the *MDISCREP* model is slightly less powerful than the *DISCREP* model.[24]

We run the following multivariate regression which is a modification of equation (6), to examine analyst forecast dispersion:

$$STD\,(Q1)_{jt} = \alpha_0 + \alpha_1 |MDISCREP_{jt}| + \alpha_2 STABLE_{jt} + \alpha_3 NONOP_{jt} + \Sigma\text{Firm} + \Sigma\text{Industry} + \Sigma\text{Year} + e_{jt} \quad (12)$$

Equation (12) differs from equation (6) in only one way; it uses the absolute value of *the magnitude of DISCREP* in lieu of *DISCREP*. Panel B of Table 8 indicates that the coefficient on |*MDISCREP*| (α_1) is positive and significant (coefficient = 0.0496, t-value = 11.11). Moreover, it has the largest t-value of any of the terms, indicating that it provides the most

[23] Unlike accuracy and dispersion, the results of revisions and market reaction cannot easily be compared using *MDISCEP* versus *DISCREP* since the *MDISCREP* formulations have one fewer explanatory variable than their *DISCREP* counterparts.
[24] Recall that it has one fewer explanatory variable.

incremental explanatory power regarding earnings forecast dispersion. Equation (12) has more total explanatory power than equation (6) [4.9% versus 4.1%], revealing that the *MDISCREP* model is more powerful than the *DISCREP* model in spite of the fact that it has one fewer explanatory variable.

We run the following multivariate regression which is a modification of equation (8), to examine market reaction to I/B/E/S earnings surprise:

$$CAR_{jt} = \alpha_0 + \alpha_1 MDISCREP_{jt} + \alpha_2 SURP_{jt} + \alpha_3 STABLE_{jt} + \alpha_5 NONOP_{jt} + \Sigma\text{Firm} + \Sigma\text{Industry} + \Sigma\text{Year} + e_{jt} \quad (13)$$

Equation (13) differs from equation (8) in two ways: (1) it uses *MDISCREP* in lieu of *DISCREP* and (2) it omits the interaction term of equation (5).[25] While equation (8) focuses on the interaction term between *DISCREP* and *SURP* and, the focal point of equation (13) is on *MDISCREP*. We expect investors to react in the direction of the analyst's inferred error which is equal to the summation of *MDISCREP* and *SURP*. Thus, we expect the coefficient on *MDISCREP* (α_1) to be positive. Consistent with our expectations, panel B of Table 8 reveals that the coefficient of *MDISCREP* is positive and significant (coefficient = 0.0096, t-value = 1.91). Equation (13) has the same total explanatory power as equation (8) [2.7% in both cases], revealing that the *MDISCREP* model is as powerful as the *DISCREP* model in spite of the fact that it has one fewer explanatory variable.

Table 8

[25] Unlike accuracy and dispersion, the formulations of revisions and market reaction cannot easily be compared using the *MIDISCREP* with the *DISCREP* measures since the multivariate *MIDISCREP* equations have one fewer explanatory variable than the corresponding *DISCREP* equations.

7. Conclusion

We examine the prevalence of, factors associated with, and consequences of data discrepancy between I/B/E/S actual EPS and analysts' inferred EPS. We find that the I/B/E/S actual EPS differs from the analyst's inferred actual EPS 35%, 45%, and 54% of the time for Q1, Q2, and Q3 respectively. Thus, the data discrepancy we discover is prevalent in the I/B/E/S earnings database. The data discrepancy we discover is systematic, being associated with at least four factors: (1) analyst, (2) firm, (3) industry, and (4) year. It is more likely to pertain to analysts who work for small brokerage houses, to firms reporting non-operating or extraordinary items, to firms followed by more analysts, to firms in the transportation and utility industries, and in more recent years. The data discrepancy introduces measurement error into studies using I/B/E/S actual EPS numbers as either their primary variables of interest or as control variables. Our investigation of the adverse consequences of the data discrepancy is confined to cases where I/B/E/S actual EPS numbers are incorporated in the primary variable of interest. We show four adverse consequences of this data discrepancy: (1) less accurate earnings forecasts by analysts; (2) smaller forecast revision coefficients by analysts; (3) more disperse earnings forecasts among analysts following a firm; (4) and lower market reactions to firms' I/B/E/S-based earnings surprises. Our results are both statistically and economically significant. The order of magnitude that we identify is large, as is evident by our univariate results for Q1 when I/B/E/S actuals differ from the analysts' inferred actuals: (1) Absolute forecast error is nearly 70% larger; (2) Forecast revision coefficients are 36% smaller; (3) Standard deviation of analyst forecasts is more than 100% larger; and (4) Capital markets react nearly 50% less to I/B/E/S-derived earnings surprises.

We provide two methods for users of I/B/E/S actual EPS data to increase the power of their tests when they examine forecast accuracy, forecast revisions, forecast dispersion, or market reaction to I/B/E/S earnings surprises. One method is to use an indicator variable for the data discrepancy, setting *DISCREP* equal to one if the I/B/E/S actual differs from the analyst's inferred actual by at least a penny. The other approach is to use a continuous variable, which is our proxy for the magnitude of the data discrepancy, which we measure as the analyst's inferred actual minus the I/B/E/S actual.

Two groups of researchers who use I/B/E/S earnings data are likely to benefit from our study: (1) researchers who do not obtain results, and (2) researchers who do obtain results. Failure to find results may be due to use of low power tests. Researchers can conduct more powerful tests by adjusting for the data discrepancy we introduce, potentially changing "no results" studies into "results" studies. The data discrepancy we introduce to the literature is less likely to change "results studies" into "no results" studies because adding power by reducing measurement error generally sharpens findings rather than overturns them. However, by employing the "fixes" we propose, researchers whose primary interest is on variables which incorporate I/B/E/S actual EPS can: (1) improve their model specifications; (2) obtain more precise estimates of their variables of interest; and (3) mitigate correlated omitted variable problems.

Our study is related to the recent study by Ljungqvist, Malloy, and Marston (2009). Both studies document that I/B/E/S data are inconsistent but the two studies differ in three major respects. First, Ljungqvist et al. find inconsistencies in the I/B/E/S recommendations data base while we demonstrate discrepancies in I/B/E/S actual earnings and analyst earnings. Second, Ljungqvist et al. assert that inconsistencies in the recommendations data base have been mitigated in recent years while we find that discrepancies in the earnings data base have become more pervasive in recent years. Third, Ljungqvist et al. maintain that researchers cannot address most of the problems they identify while we offer researchers two methods for addressing all the problems we identify.

APPENDIX 1: Variable Definitions of Terms Used in the Primary Analysis

Variable		Definition
Accuracy (FY)$_{ijt}$	=	$\left\lvert \frac{EPS(FY)_{jt} - AF(FY)_{ijt}}{price_{jt-1}} \right\rvert$
Accuracy (Q1)$_{ijt}$	=	$\left\lvert \frac{EPS(Q1)_{jt} - AF(Q1)_{ijt})}{price_{jt-1}} \right\rvert$
AF (FY)$_{ijt}$	=	Analyst *i*'s forecast of fiscal year *t* EPS according to I/B/E/S
AF (Q1)$_{ijt}$	=	Analyst *i*'s forecast of Q1 EPS for year *t* according to I/B/E/S
BSIZE$_{ijt}$	=	Analyst *i*'s brokerage size, calculated as the number of analysts employed by the brokerage house of analyst *i* following firm *j* in year *t* minus the minimum number of analysts employed by brokerage houses of all analysts following firm *j* in year *t*, with this difference scaled by the range of brokerage house sizes for all analysts following firm *j* in year *t*
CAR$_{jt}$	=	Two-day (-1,0) cumulative return minus cumulative value-weighted market return related to the announcement of Q1 earnings for firm *j* in year *t* where 0 is the Q1 earnings announcement day
DISCREP$_{ijt}$	=	1 when analyst *i* following firm *j* has an annual forecast for year *t* made after the release of Q1 EPS in year *t* that exceeds the sum of the I/B/E/S Q1 actual EPS and the analyst's EPS forecast of the remainder of year *t* by at least one penny
DISCREP$_{jt}$	=	1 when at least one analyst following firm *j* has an annual forecast for year *t* made after the release of Q1 EPS in year *t* that exceeds the sum of the I/B/E/S Q1 actual EPS and the analyst's EPS forecast of the remainder of year *t* by at least one penny
EPS (FY)$_{jt}$	=	Actual annual EPS for fiscal year *t* according to I/B/E/S
EPS (Q1)$_{jt}$	=	Actual Q1 EPS for fiscal year *t* according to I/B/E/S
FEXP$_{ijt}$	=	Analyst *i*'s firm experience, calculated as the number of prior forecasting years for analyst *i* following firm *j* in year *t* minus the minimum number of prior forecasting years for all analysts following firm *j* in year *t*, with this difference scaled by the range of prior forecasting years for all analysts following firm *j* in year *t*
FREQ$_{ijt}$	=	Analyst *i*'s forecast frequency, calculated as the number of firm *j* forecasts made by analyst *i* following firm *j* in year *t* minus the minimum

		number of firm *j* forecasts for all analysts following firm *j* in year *t*, with this difference scaled by the range of number of firm *j* forecasts issued by all analysts following firm *j* in year *t*
$HORIZON_{ijt}$	=	Analyst *i*'s horizon, calculated as the number of days from the year *t* forecast date (following the release of Q1 earnings) to the earnings announcement date for analyst following firm *j* in year *t* minus the minimum forecast horizon for all analysts who follow firm *j* in year *t*, with this difference scaled by the range of forecast horizons for all analysts following firm *j* in year *t*
$MDISCREP_{ijt}$	=	Analyst inferred actual – IBES actual, where the analyst inferred actual is the difference between analyst *i*'s forecast for the fiscal year and the sum of the analyst's forecasts for Q2, Q3, and Q4 of the same year
$NFIRMS_{ijt}$	=	The number of firms followed, calculated as the total number of firms followed by analyst *i* following firm *j* in year *t* minus the minimum number of firms followed by all analysts covering firm *j* in year *t*, with this difference scaled by the range of the number of firms followed by all analysts covering firm *j* in year *t*
$NONOP_{jt}$	=	1 when the absolute value of the difference between Q1 EPS Compustat operating earnings and Compustat earnings after extraordinary items for year *t* exceeds one penny for firm *j* in year *t*
$Price_{jt-1}$	=	Stock price as of the end of fiscal year *t-1* according to CRSP
REV_9M_{ijt}	=	The difference between analyst *i*'s first estimate of Q2 through Q4 issued after the release of Q1 earnings and the same analyst's last estimate of Q2 through Q4 issued before the release of Q1 earnings, scaled by stock price as of the end of fiscal year *t-1* according to CRSP
$STABLE_{jt}$	=	I/B/E/S stability measure, defined by I/B/E/S as a gauge of annual EPS growth consistency over the past 5 years
STD_{jt}	=	Standard deviation of analysts' EPS estimates for firm j, for a given quarter or year t, for those firms with >1 analyst following in our sample
$SURP_{jt}$	=	$\frac{EPS(Q1)_{jt} - AF(Q1)_{jt}}{price_{jt-1}}$

APPENDIX 2: Calculation of the *DISCREP* variable

Assume:

EPS (Q1) = Actual Q1 EPS according to I/B/E/S

AF (FY/Q1) = Analyst's forecast of annual EPS according to I/B/E/S issued after the quarter 1 earnings announcement

AF(Q2-Q4/Q1)= Analyst's forecast of EPS for the remainder of the fiscal year issued after the quarter 1 earnings announcement

We define the *Analyst's Inferred EPS* (*Q1) = AF*(*FY/Q1*) – *AF*(*Q2-Q4/Q1)*

DISCREP = 1 if:

|Analyst Inferred EPS (*Q1) – EPS (Q1)| >* 0.01

and

DISCREP = 0 if:

|Analyst Inferred EPS (*Q1) – EPS*(*Q1*)*| ≤* 0.01

References

Ball, R. and P. Brown. 1968. An empirical evaluation of accounting income numbers. *Journal of Accounting Research* 6 (2): 159-178.

Barron, O., D. Byard, and O. Kim. 2002. Changes in analysts' information around earnings announcements. *The Accounting Review* 77 (4): 821-846.

Beaver, W. 1998. Financial Reporting: An Accounting Revolution, Third Edition, Prentice Hall, Englewood Cliffs, New Jersey.

Beaver, W., R. Clarke, and W. Wright. 1979. The association between unsystematic security returns and the magnitude of earnings forecast errors. *Journal of Accounting Research* 17 (2): 316-340.

Bhattacharya, N., E. Black, T. Christensen, and C. Larson. 2003. Assessing the relative informativeness and permanence of pro forma earnings and GAAP operating earnings. *Journal of Accounting and Economics* 36 (1-3): 285-319.

Bradshaw, M. 2003. A discussion of 'Assessing the relative informativeness and permanence of pro forma earnings and GAAP operating earnings.' *Journal of Accounting and Economics* 36 (1-3): 321-335.

Bradshaw, M. and R. Sloan. 2002. GAAP versus the street: An empirical assessment of two alternative definitions of earnings. *Journal of Accounting Research* 40 (1): 41-66.

Brown, L. (editor). 2007. Thomson Financial Research Bibliography. First Edition. Thomson. New York.

Brown, L. and K. Kim. 1991. Timely aggregate analyst forecasts as better proxies for market earnings expectations. *Journal of Accounting Research* 29 (2): 382-385.

Brown, L. and K. Sivakumar. 2003. Comparing the value relevance of two operating income measures. *Review of Accounting Studies* 8 (4): 561-572.

Brown, L. and M. Rozeff. 1979. Adaptive expectations, time-series models, and analyst forecast revision. *Journal of Accounting Research* 17 (2): 341-351.

Brown, P. and J. Kennelly. 1972. The informational content of quarterly earnings: An extension and some further evidence. *Journal of Business* 45 (3): 403-415.

Clement, M. 1999. Analyst forecast accuracy: Do ability, resources and portfolio complexity matter? *Journal of Accounting and Economics* 27 (3): 285-303.

Clement, M. and S. Tse. 2005. Financial analyst characteristics and herding behavior in forecasting. *Journal of Finance* 60 (1): 307-341.

Doyle, J., R. Lundholm, and M. Soliman. 2003. The predictive value of expenses excluded from pro forma earnings. *Review of Accounting Studies* 8 (2-3): 145-174.

Elliott, J. and J. Hanna. 1996. Repeated accounting write-offs and the information content of earnings. *Journal of Accounting Research* 34 (Supplement): 135-155.

Ertimur, Y., J. Sunder, and S. Sunder. 2007. Measure for measure: The relation between forecast accuracy and recommendation profitability of analysts. *Journal of Accounting Research* 45 (3): 567-606.

Feng, M. and S. McVay. 2010. Analysts' incentives to overweight management guidance when revising their short-term earnings forecasts. *The Accounting Review* 85 (5): 1617-1646.

Foster, G. 1977. Quarterly accounting data: Time series properties and predictive-ability results. *The Accounting Review* 52 (1): 1-21.

Hamilton, L. 1992. Regression with Graphics. Wadsworth, Inc.

Hugon, A. and V. Muslu. 2010. Market demand for conservative analysts. *Journal of Accounting and Economics* 50 (1): 42-57.

Jacob, J., T. Lys, and M. Neale. 1999. Expertise in forecasting performance of security analysts. *Journal of Accounting and Economics* 28 (1): 51-82.

Ljungqvist, A., C. Malloy, and F. Marston. 2009. Rewriting history. *Journal of Finance* 64 (4): 1935-1960.

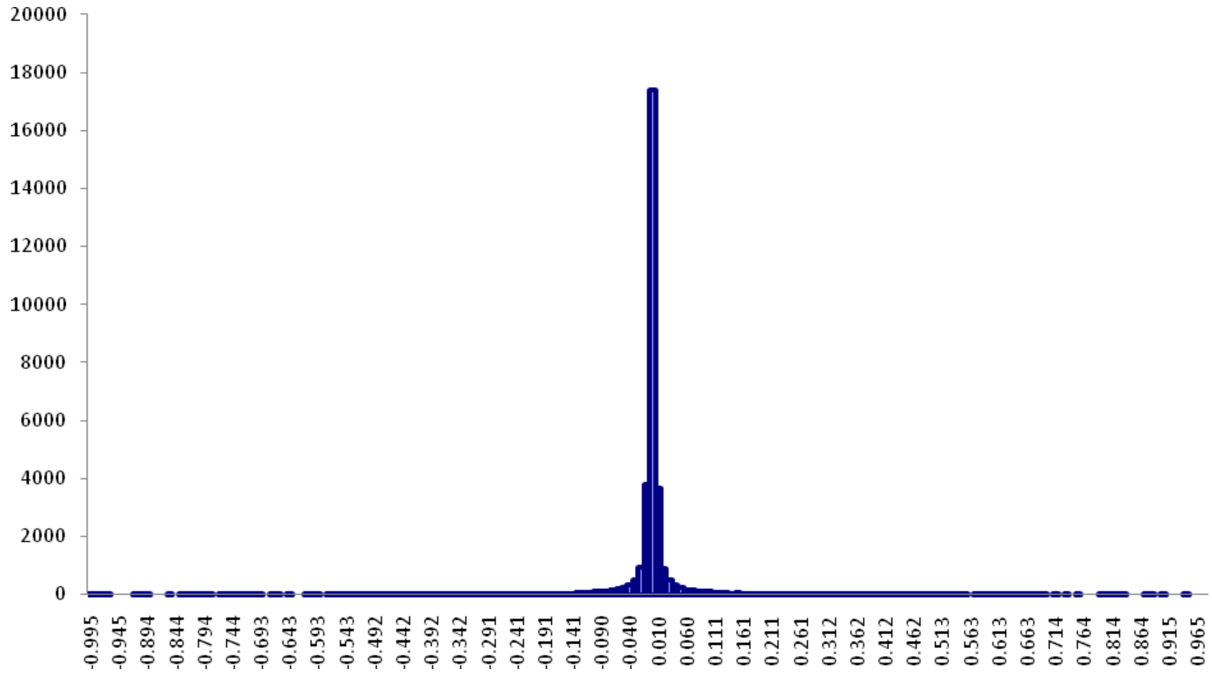
Mikhail, M., B. Walther, and R. Willis. 1997. Do security analysts improve their performance with experience? *Journal of Accounting Research* 35 (Supplement): 131-157.

O'Brien, P. 1990. Forecast accuracy of individual analysts in nine industries. *Journal of Accounting Research* 28 (2): 286-304.

Payne, J., and W. Thomas. 2003. The implications of using stock-split adjusted I/B/E/S data in empirical research. *The Accounting Review* 78: 1049-1067.

Ramnath, S., S. Rock, and P. Shane. 2008. The financial analyst forecasting literature: taxonomy with suggestions for future research. *International Journal of Forecasting* 24 (1): 34-75.

FIGURE 1
Distribution of Differences between I/B/E/S Q1 Actual EPS and Analysts' Inferred Q1 EPS



This Figure presents the distribution of the difference between the analysts' inferred Q1 EPS and the I/B/E/S Q1 actual for the analyst-firm-years in our sample. Variable definitions are in the Appendix.

TABLE 1
Sample Selection Procedure for the Main Analysis

Criteria	Analyst-firm-years	Firm-years
Firm-years:		
U.S. firm-years with FY_{t-1}, $Q1_t$, and $Q2_t$ reporting dates available from I/B/E/S for 1996 to 2008		56,366
Keep: Firm-years with I/B/E/S actual $Q1_t$ and FY_t earnings		47,615
Keep: Firm-years with year t-1 prices available from CRSP		46,037
Analyst-firm-years:		
Analyst-firm-years with $Q2_t$, $Q3_t$, $Q4_t$, and FY_t EPS forecasts issued on the same day after the release of $Q1_t$ earnings	83,258	33,034
Keep: Analyst-firm-years with $Q1_t$ EPS forecasts issued on any day after the release of FY_{t-1} earnings but prior to the release of $Q1_t$ earnings (Table 3)	31,818	18,483
Analyst-firm-years with 1 analyst following in our sample	11,375	11,375
Analyst firm-years with >1 analyst following in our sample	20,443	7,108
Sub-samples:		
Analyst-firm-years with FY_t EPS forecasts issued after the release of FY_{t-1} earnings but prior to the release of $Q1_t$ earnings (Table 3)	23,921	14,771
Analyst-firm-years with $Q2_t$, $Q3_t$, and $Q4_t$ EPS forecasts issued after the release of FY_{t-1} earnings but prior to the release of $Q1_t$ earnings (Table 4)	20,968	13,955
Analyst-firm-years with FY_t EPS forecasts issued after the release of FY_{t-1} earnings but prior to the release of $Q1_t$ earnings, with > 1 analyst following (Table 5)	15,824	5,078
Firm-years with non-missing CAR_t (Table 6)		18,465

This table summarizes the procedure used to select the sample for Tables 3-6. Variable definitions are in Appendix 1.

TABLE 2
Estimation of Model with Analyst, Firm, Industry, and Year Fixed Effects

Panel A Logit Model	Analysts (1)	Firms (2)	I/B/E/S Industries (3)	Years (4)	Full Model[1] (5)
N	3,706	4,317	11	13	28,394
Log-likelihood	-8,466.0	-11,857.3	-18,162.5	-18,069.8	-3,436.9
Prob > χ^2	0.00	0.00	0.00	0.00	0.00

[1] Sample sizes and full-model log-likelihood are reported.

Panel B Analyst Effects	Intercept (1)	$FEXP_{ijt}$ (2)	$FREQ_{ijt}$ (3)	$NFIRMS_{ijt}$ (4)	$BSIZE_{ijt}$ (5)	$HORIZON_{ijt}$ (6)	% Concordant	Number of analyst-firm-years
Pr($DISCREP_{ijt}$)=1	0.6448*** (0.0520)	0.0215 (0.0388)	0.0295 (0.0532)	-0.0240 (0.0454)	-0.2001*** (0.0465)	-0.0253 (0.0375)	50.0	20,881

Panel C Firm Effects	ANF=1 (1)	ANF=2 (2)	ANF=3 (3)	ANF=4 (4)	ANF≥5 (5)
# analyst-firm-years	11,375	8,290	4,512	2,732	4,909
% $DISCREP_{ijt}$	0.327	0.337	0.351	0.366	0.389

TABLE 2 (continued)

Panel D Firm Effects	Basic (1)	Capital (2)	ConsDur (3)	ConsNonDur (4)	ConsSvc (5)	Energy (6)	Finance (7)	Health (8)	Technol (9)	Transp (10)	Utility (11)	Total
Total # analyst-firm-years	1,863	1,348	911	618	2,938	3,645	3,686	1,661	2,927	927	725	21,249
Correlation	0.019	0.073***	0.063**	0.042	0.101***	0.051***	0.108***	0.056**	0.063***	0.153***	0.019	0.068***
($DISCREP_{ijt}$, $NONOP_{jt}$)	(0.936)	(0.008)	(0.059)	(0.293)	(<0.0001)	(0.002)	(<0.0001)	(0.022)	(0.001)	(<0.001)	(0.611)	(<0.001)

Panel E Industry Effects	Basic (1)	Capital (2)	ConsDur (3)	ConsNonDur (4)	ConsSvc (5)	Energy (6)	Finance (7)	Health (8)	Technol (9)	Transp (10)	Utility (11)	Total
# analyst-firm-years	2,829	2,044	1,343	1,019	4,502	4,706	5,248	2,512	5,026	1,368	1,214	31,811
% $DISCREP_{ijt}$	33.4	27.3	32.7	29.2	36.7	38.9	32.3	32.6	33.3	42.7	42.1	34.6

TABLE 2 (continued)

Panel F Year Effects	Basic (1)	Capital (2)	ConsDur (3)	ConsNonDur (4)	ConsSvc (5)	Energy (6)	Finance (7)	Health (8)	Technol (9)	Transp (10)	Utility (11)	Total
a	0.894***	1.292***	-1.092***	-1.246***	-0.845***	-0.907***	-1.188***	-1.128***	-1.043***	-0.569***	-0.564***	0.992***
b	0.028***	0.041	0.049*	0.047***	0.038***	0.053***	0.050***	0.050***	0.051***	0.032**	0.032**	0.045***
% Concordant	49.3	50.7	51.4	51.3	50.0	51.7	51.0	51.5	51.4	48.7	49.6	51.0
# analyst-firm-years	2,829	2,044	1,343	1,019	4,502	4,706	5,248	2,512	5,026	1,368	1,214	31,811

This table investigates the effects of analysts, firms, industries, and years on the *DISCREP* indicator variable. Panel A estimates the following equation using a logit specification:

$$DISCREP_{ijkt} = \mu_i + \delta_j + \lambda_k + \gamma_t + \varepsilon_{ijkt}$$

In Panel A, observations in which the analyst or firm appears only once in the full sample are deleted in order to estimate the fixed effects models.

Panel B presents results based on estimating the following equation using a logit specification for the sub-sample of analyst-firm-years for which the analyst variables could be estimated (standard errors are in parentheses):

$$Pr(DISCREP_{ijt}=1) = \alpha_0 + \alpha_1 FEXP_{ijt} + \alpha_2 FREQ_{ijt} + \alpha_3 NFIRMS_{ijt} + \alpha_4 BSIZE_{ijt} + \alpha_5 HORIZON_{ijt} + e_{ijt}$$

Panel C provides the number and percent of analyst-firm-years for *DISCREP* for one, two, three, four, and five or more analysts following the firm.

Panel D presents the correlation of *DISCREP* with *NONOP* for analyst-firm-years that could be matched with Compustat for 11 I/B/E/S industries: Basic Industries (Basic), Capital Goods (Capital), Consumer Durables (ConsDur), Consumer Non-Durables (ConsNonDur), Consumer Services (ConsSvc), Energy (Energy), Finance (Finance), Health Care (Health), Technology (Technol), Transportation (Transp), and Public Utilities (Utility). Standard errors are in parentheses.

Panel E provides the number and percent of analyst-firm-years for *DISCREP* for each I/B/E/S industry and for the full sample.

Panel F presents results of estimating the following specification for each I/B/E/S industry and for the full sample:

$Pr(DISCREP_{ijt}=1) = a + b*Year_t + e_{ijt}$

Results are pooled. ***, **, and * denote significance at the 1%, 5%, and 10% levels, based on one-tailed tests, respectively. Variable definitions are in the Appendix.

TABLE 3
Analysts' Earnings Forecast Accuracy

	All observations (1)	DISCREP = 0 (2)	DISCREP = 1 (3)	(3) – (2)
A. Analyses without control variables				
Accuracy (Q1)$_{ijt}$	0.0048	0.0039	0.0066	0.0027*** (18.19)
N	31,818	20,812	11,006	
Accuracy (FY)$_{ijt}$	0.0445	0.0379	0.0567	0.0188*** (12.18)
N	23,921	15,508	8,413	

TABLE 3 (continued)

B. Analyses with control variables

Dependent variable is:		$Accuracy(Q1)_{ijt}$	$Accuracy(FY)_{ijt}$
Intercept	?	0.0005	0.0114***
		(1.37)	(3.57)
$DISCREP_{ijt}$	+	0.0004***	0.0023**
		(3.12)	(2.01)
$FEXP_{ijt}$	-	0.0006**	0.0068***
		(2.42)	(3.12)
$FREQ_{ijt}$	-	-0.0004	0.0003
		(-1.54)	(0.12)
$NFIRMS_{ijt}$	+	0.0005*	0.0017
		(1.68)	(0.63)
$BSIZE_{ijt}$	-	-0.0001	0.0004
		(-0.14)	(0.11)
$HORIZON_{ijt}$	+	0.0004*	0.0034**
		(1.94)	(1.93)
$NONOP_{jt}$	+	0.0004***	0.00005
		(3.27)	(0.04)
Analyst Effects		Yes	Yes
Firm Effects		Yes	Yes
Industry Effects		Yes	Yes
Year Effects		Yes	Yes
Adj. R^2		0.010	0.017
Number of Analyst-firm-years		13,836	11,842

TABLE 3 (continued)

This table estimates the accuracy of analysts' EPS estimates. *Accuracy* is defined as the absolute value of the difference between actual Q1 (FY) earnings and the analyst's last EPS estimate of Q1 (FY) preceding the release of Q1 EPS, scaled by lagged stock price. Both actual earnings and analysts' earnings forecasts are obtained from I/B/E/S. Stock price is obtained from CRSP. *DISCREP* equals zero when analyst i following firm j has an annual forecast for year t made after the release of Q1 EPS in year t that is within one penny of the sum of the I/B/E/S Q1 actual and the analyst's forecast of the remainder of year t.. *DISCREP* equals one when analyst i following firm j has an annual forecast for year t made after the release of Q1 EPS in year t that is not within one penny of the sum of the I/B/E/S Q1 actual and the analyst's forecast of the remainder of year. Panel A presents univariate results. Panel B presents multivariate results which control for analyst characteristics, *NONOP*, analyst effects, firm effects, industry effects, and year effects using the following specification:

$$\begin{aligned} Accuracy_{ijt} = \quad & \alpha_0 + \alpha_1 DISCREP_{ijt} + \alpha_2 FEXP_{ijt} + \alpha_3\, FREQ_{ijt} + \alpha_4 NFIRMS_{ijt} + \alpha_5 BSIZE_{ijt} + \alpha_6 HORIZON_{ijt} \\ & + \alpha_7 NONOP_{jt} + \Sigma Analyst + \Sigma Firm + \Sigma Industry + \Sigma Year + e_{ijt} \end{aligned}$$

Results are pooled. T-statistics are in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, based on one-tailed tests, respectively. All variable definitions are in Appendix 1.

TABLE 4
Analysts' Earnings Forecast Revisions

	All observations (1)	DISCREP = 0 (2)	DISCREP = 1 (3)	(3) – (2)
A. Analyses without control variables				
b	1.0946***	1.3667***	0.8656***	-0.5011***
	(45.32)	(42.11)	(23.02)	(-10.35)
Adj. R^2	0.089	0.115	0.068	
N	20,968	13,666	7,302	

TABLE 4 (continued)

B. Analyses with control variables

Dependent variable is:		REV_9M_{ijt}
Intercept	?	-0.0012
		(-1.29)
$DISCREP_{ijt}$	?	0.0006*
		(1.65)
$SURP_{ijt}$	+	0.9205***
		(26.52)
$DISCREP_{ijt}$ * $SURP_{ijt}$	-	-0.2505***
		(-5.08)
$NONOP_{jt}$	-	-0.0001
		(-0.25)
Analyst Effects		Yes
Firm Effects		Yes
Industry Effects		Yes
Year Effects		Yes
Adj. R^2		0.082
Number of Analyst-firm-years		13,011

TABLE 4 (continued)

This table estimates the revision of analysts' EPS forecasts. The revision of the analyst's EPS estimate for Q2 through Q4 (*REV_9M*) is defined as the difference between the analyst's first 9M estimate after the release of Q1 earnings and the same analyst's last 9M estimate before the release of Q1 earnings, scaled by lagged stock price, where the 9M estimate is the summation of the estimates of Q2, Q3, and Q4. The I/B/E/S earnings surprise (*SURP*) is defined as the difference between the I/B/E/S Q1 actual earnings and the same analyst's last Q1 estimate before the release of Q1 earnings, scaled by lagged stock price *DISCREP* equals zero when analyst i following firm j has an annual forecast for year t made after the release of Q1 EPS in year t that is within one penny of the sum of the I/B/E/S Q1 actual and the analyst's forecast of the remainder of year t. *DISCREP* equals one when analyst i following firm j has an annual forecast for year t made after the release of Q1 EPS in year t that is not within one penny of the sum of the I/B/E/S Q1 actual and the analyst's forecast of the remainder of year.Panel A uses the following specification:

$$REV\_9M_{ijt} = \quad a + b*SURP_{ijt} + e_{ijt}$$

Results for the intercepts are excluded for simplicity. The final column compares the slope coefficients in columns 2 and 3.

Panel B presents results which control for *NONOP*, and analyst, firm, industry, and year effects:

$$REV\_9M_{ijt} = \quad \alpha_0 + \alpha_1 DISCREP_{ijt} + \alpha_2 SURP_{ijt} + \alpha_3\, DISCREP_{ijt} * SURP_{ij} + \alpha_4 NONOP_{jt}$$

$$+ \Sigma Analyst + \Sigma Firm + \Sigma Industry + \Sigma Year + e_{ijt}$$

Results are pooled. T-statistics are in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, based on one-tailed tests, respectively. All variable definitions are in Appendix 1.

TABLE 5
Dispersion of Analysts' Earnings Forecasts

	All observations (1)	DISCREP = 0 (2)	DISCREP = 1 (3)	(3) – (2)
A. Analyses without control variables				
STD $(Q1)_{ijt}$	0.0440	0.0269	0.0556	0.0287***
				(10.57)
N	7,108	2,868	4,240	
STD $(FY)_{ijt}$	0.1803	0.1100	0.2204	0.1104***
				(10.26)
N	5,078	1,842	3,236	

TABLE 5 (continued)

B. Analyses with control variables

Dependent variable is:		$STD(Q1)_{jt}$	$STD(FY)_{jt}$
Intercept		0.0186***	0.0817***
		(5.56)	(6.36)
$DISCREP_{jt}$	+	0.0043***	0.0140***
		(3.85)	(3.55)
$STABLE_{jt}$	-	-0.00004***	0.00002
		(-2.56)	(0.32)
$NONOP_{jt}$	+	0.0016*	-0.0050
		(1.48)	(-1.30)
Firm Effects		Yes	Yes
Industry Effects		Yes	Yes
Year Effects		Yes	Yes
Adj. R^2		0.041	0.075
Number of Firm-years		13,237	10,869

This table estimates the dispersion of analysts' EPS estimates. Dispersion is defined as the average of the standard deviation of analysts' EPS estimates for a given interval for those firms with >1 analyst following in our sample. *DISCREP* equals zero when analyst i following firm j has an annual forecast for year t made after the release of Q1 EPS in year t that is within one penny of the sum of the I/B/E/S Q1 actual and the analyst's forecast of the remainder of year t. *DISCREP* equals one when analyst i following firm j has an annual forecast for year t made after the release of Q1 EPS in year t that is not within one penny of the sum of the I/B/E/S Q1 actual and the analyst's forecast of the remainder of year. Panel A presents univariate results without control variables. Panel B presents results which control for *STABLE*, *NONOP*, and firm, industry, and year effects, using the following specification:

$$STD\ (Q1)_{jt} = \alpha_0 + \alpha_1 DISCREP_{jt} + \alpha_2 STABLE_{jt} + \alpha_3 NONOP_{jt} + \Sigma Firm + \Sigma Industry + \Sigma Year + e_{jt}$$

Results are pooled. T-statistics are in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, based on one-tailed tests, respectively. All variable definitions are in Appendix 1.

TABLE 6
Market Reaction to I/B/E/S Earnings Surprises

	All observations (1)	DISCREP = 0 (2)	DISCREP = 1 (3)	(3) – (2)
A. Analyses without control variables				
B	0.9761***	1.3640***	0.7165***	-0.6475***
	(17.78)	(15.79)	(10.06)	(-5.78)
Adj. R^2	0.017	0.023	0.013	
N	18,465	10,511	7,954	

B. Analyses with control variables

Dependent variable is:		CAR_{jt}
Intercept	?	0.0017
		(0.48)
$DISCREP_{jt}$	?	-0.0002
		(-0.18)
$SURP_{jt}$	+	2.1154***
		(11.95)
$DISCREP_{jt}$ * $SURP_{jt}$	-	-0.6205***
		(-2.75)
$STABLE_{jt}$	+	0.00003*
		(1.50)
$NONOP_{jt}$	-	0.0010
		(0.80)
Firm Effects		Yes
Industry Effects		Yes
Year Effects		Yes
Adj. R^2		0.027
Number of Firm-years		8,810

TABLE 6 (continued)

This table estimates the stock market reaction to first quarter I/B/E/S earnings surprise. *CAR* represents two-day (-1, 0) abnormal returns where day 0 is the release of Q1 earnings. *DISCREP* equals zero when analyst i following firm j has an annual forecast for year t made after the release of Q1 EPS in year t that is within one penny of the sum of the I/B/E/S Q1 actual and the analyst's forecast of the remainder of year t.. *DISCREP* equals one when analyst i following firm j has an annual forecast for year t made after the release of Q1 EPS in year t that is not within one penny of the sum of the I/B/E/S Q1 actual and the analyst's forecast of the remainder of year. Panel A uses the following specification:

$$CAR_{jt} = a + b*SURP_{jt} + e_{jt}$$

Results for intercepts are excluded for simplicity. The final column compares the slope coefficients in columns 2 and 3.

Panel B presents results which control for *STABLE*, *NONOP*, and firm, industry, and year effects, using the following specification:

$$CAR_{jt} = \alpha_0 + \alpha_1 DISCREP_{jt} + \alpha_2 SURP_{jt} + \alpha_3\ DISCREP_{jt} * DISCREP_{jt}*SURP_{jt} + \alpha_4 STABLE_{jt} + \alpha_5 NONOP_{jt}$$
$$+ \Sigma Firm + \Sigma Industry + \Sigma Year + e_j$$

Results are pooled. T-statistics are in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, based on one-tailed tests, respectively. All variable definitions are in Appendix 1.

TABLE 7

Analyses of Second and Third Interim Quarters

Panel A: Dependent variable is:		Accuracy(Q2)$_{ijt}$ (1)	Accuracy(Q3)$_{ijt}$ (2)	REV_6M$_{ijt}$ (3)	REV_3M$_{ijt}$ (4)
Intercept	?	0.0009**	0.0010***	-0.0017**	-0.0016***
		(1.98)	(3.39)	(-2.40)	(-6.17)
DISCREP$_{ijt}$	+/?	0.0002	0.0007***	-0.0001	0.0001
		(1.13)	(4.26)	(-0.28)	(0.57)
SURP$_{ijt}$	+			0.6463***	0.3259***
				(43.58)	(59.47)
DISCREP$_{jt}$ * SURP$_{ijt}$	-			-0.0993***	-0.0609***
				(-3.71)	(-4.93)
FEXP$_{ijt}$	-	0.0006**	0.0014***		
		(2.07)	(7.00)		
FREQ$_{ijt}$	-	0.0008**	0.0004**		
		(2.14)	(2.02)		
NFIRMS$_{ijt}$	+	0.0004	0.0002		
		(1.06)	(0.97)		
BSIZE$_{ijt}$	-	0.00002	0.0004		
		(0.03)	(1.05)		
HORIZON$_{ijt}$	+	0.0006**	0.0003**		
		(2.43)	(2.03)		
NONOP$_{jt}$	+/-	0.0012***	0.0009***	0.0003	-0.0002***
		(7.28)	(8.81)	(1.12)	(-2.56)
Analyst Effects		Yes	Yes	Yes	Yes
Firm Effects		Yes	Yes	Yes	Yes
Industry Effects		Yes	Yes	Yes	Yes
Year Effects		Yes	Yes	Yes	Yes
Adj. R^2		0.014	0.014	0.115	0.107
Number of Analyst-firm-years		15,576	33,829	21,180	38,570

TABLE 7 (continued)

Panel B: Dependent variable is:		$STD(Q2)_{jt}$ (1)	$STD(Q3)_{jt}$ (2)	$CAR(Q2)_{jt}$ (3)	$CAR(Q3)_{jt}$ (4)
Intercept	?	0.0200***	0.0244***	-0.0045	-0.0045
		(5.66)	(10.37)	(-0.83)	(-0.84)
$DISCREP_{jt}$	+/?	0.0052***	0.0091***	-0.0008	-0.0007
		(4.08)	(6.61)	(-0.53)	(-0.48)
$SURP_{jt}$	+			1.5153***	0.8244***
				(11.44)	(8.10)
$DISCREP_{jt}$ * $SURP_{j}$	-			-0.3894**	-0.5476***
				(-1.86)	(-2.95)
$STABLE_{jt}$	-/+	0.00004***	0.00004***	-0.00001	0.00002
		(2.50)	(3.10)	(-0.29)	(0.70)
$NONOP_{jt}$	+/-	0.0059***	0.0079***	0.0005	0.0003
		(5.47)	(9.32)	(0.36)	(0.20)
Firm Effects		Yes	Yes	Yes	Yes
Industry Effects		Yes	Yes	Yes	Yes
Year Effects		Yes	Yes	Yes	Yes
Adj. R^2		0.066	0.039	0.019	0.005
Number of Firm-years		14,542	41,742	9,041	11,170

TABLE 7 (continued)

This table repeats the panel B analyses in Tables 3 to 6 across two time intervals subsequent to the first interim quarter: after the release of Q2 but before the release of Q3 EPS; and after the release of Q3 but before the release of Q4 EPS. Panel A uses analyst-firm-year data and investigates accuracy and revisions. Panel B uses firm-year data and investigates dispersion and market reactions to earnings surprises.

Results are pooled. T-statistics are in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, based on one-tailed tests, respectively. All variable definitions are in the Appendix except as discussed next. *Accuracy* (Q2) [Q3] is the absolute value of actual Q2 [Q3] EPS according to I/B/E/S minus the analyst's forecast of Q2 [Q3] EPS according to I/B/E/S, scaled by stock price as of the end of the previous fiscal year according to CRSP. *Rev_6M* is the difference between an analyst's first estimate of Q3 through Q4 issued after the release of Q2 earnings and the same analyst's last estimate of Q3 through Q4 issued before the release of Q2 earnings, scaled by lagged stock price from CRSP. *Rev_3M* is the difference between an analyst's first estimate of Q4 issued after the release of Q3 earnings and the same analyst's last estimate of Q4 issued before the release of Q3 earnings, scaled by lagged stock price from CRSP. *CAR* (Q2) [Q3] is the two-day (-1,0) cumulative return minus the cumulative value-weighted market return related to the announcement of Q2 [Q3] earnings for firm j in year t. *DISCREP* in the Q2 [Q3] analyst-firm-year analysis in panel A equals one when analyst i following firm j in industry k has an annual forecast for year t made after the release of Q2 [Q3] EPS in year t that is not within one penny of the analyst's forecast of the remainder of year t plus the I/B/E/S Q1 and Q2 [Q1, Q2, and Q3] actuals, and zero otherwise. *DISCREP* in panel B denotes firm-years for which at least one analyst following the firm has an annual forecast made after the release of Q2 [Q3] EPS that differs from the analyst's forecast of the remainder of the year plus the I/B/E/S Q1 and Q2 [Q1, Q2, and Q3] actuals by one penny or more. *HORIZON* in the Accuracy (Q2) [Q3] analysis is calculated as the number of days from the year t forecast date (following the release of Q2 [Q3] earnings) to the earnings announcement date for analyst i following firm j in year t minus the minimum forecast horizon for all analysts who follow firm j in year t, with this difference scaled by the range of forecast horizons for all analysts following firm j in year t. *NONOP* in the Q2 [Q3] analyses equals 1 when the absolute value of the difference between Q2 [Q3] EPS Compustat operating earnings and Compustat earnings after extraordinary items for year t exceeds one penny. *SURP* in the consecutive REV and CAR analyses is actual Q2 [Q3] EPS according to I/B/E/S minus the analyst's forecast of Q2 [Q3] EPS according to I/B/E/S scaled by stock price as of the end of the previous fiscal year according to CRSP.

TABLE 8
Analyses Using a Continuous Measure of Data Discrepancy

Panel A: Dependent variable is:		$Accuracy(Q1)_{ijt}$	REV_9M_{ijt}
Intercept	?	0.0007*	-0.0009
		(1.94)	(-0.95)
$\lvert MDISCREP_{ijt}\rvert$	+	0.0001***	
		(4.11)	
$MDISCREP_{ijt}$	+		0.0003**
			(2.11)
$SURP_{ijt}$	+		0.7960***
			(32.20)
$FEXP_{ijt}$	-	0.0006**	
		(2.36)	
$FREQ_{ijt}$	-	-0.0005	
		(-1.55)	
$NFIRMS_{ijt}$	+	0.0005*	
		(1.65)	
$BSIZE_{ijt}$	-	-0.00002	
		(-0.05)	
$HORIZON_{ijt}$	+	0.0004*	
		(1.94)	
$NONOP_{jt}$	+	0.0004***	-0.0001
		(3.50)	(-0.19)
Analyst Effects		Yes	Yes
Firm Effects		Yes	Yes
Industry Effects		Yes	Yes
Year Effects		Yes	Yes
Adj. R^2		0.011	0.081
Number of Analyst-firm-years		13,836	13,011

TABLE 8 (continued)

Panel B: Dependent variable is:		STD (Q1)$_{jt}$	CAR$_{jt}$
Intercept	?	0.0201*** (6.09)	0.0017 (0.47)
$\lvert$MDISCREP$_{jt}\rvert$	+	0.0496*** (11.11)	
MDISCREP$_{jt}$	+		0.0096** (1.91)
SURP$_{jt}$	+		1.7833*** (15.79)
STABLE$_{jt}$	-/+	-0.00004*** (-2.61)	0.00003* (1.46)
NONOP$_{jt}$	+/-	0.0016* (1.54)	0.0009 (0.74)
Firm Effects		Yes	Yes
Industry Effects		Yes	Yes
Year Effects		Yes	Yes
Adj. R^2		0.049	0.027
Number of Firm-years		13,237	8,810

This table repeats the panel B analyses in Tables 3 to 6 using a continuous measure of data discrepancy, the difference between the analyst's inferred Q1 actual and the I/B/E/S Q1 actual. Panel A uses analyst-firm-year data and investigates accuracy and revisions. This continuous measure is denoted as MDISCREP to designate the magnitude of the discrepancy. Panel B uses firm-year data and investigates dispersion and market reactions to I/B/E/S earnings surprises. Results are pooled. T-statistics are in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, based on one-tailed tests, respectively. All variable definitions are in Appendix 1.